

A Skew Normal Stochastic Frontier Model for Panel Data

Roberto Colombi

Abstract We introduce a stochastic frontier model for longitudinal data where a subject random effect coexists with a time independent random inefficiency component and with a time dependent random inefficiency component. The role of the closed skew normal distribution in this kind of modeling is stressed.

Key words: Longitudinal data analysis, Statistical Modeling, Variance Components Models

1 Introduction

A stochastic frontier defines a random upper bound of a response variable called *output* as a function of variables called *inputs* plus an idiosyncratic random error. The observable *output* is given by the stochastic frontier minus a non negative random variable called subject random inefficiency component. Stochastic frontiers have been mainly used as statistical models for production functions (Kumbhakar and Lovell, 2000) but we believe that they are random effects models of wider applicability. In this paper a three random component stochastic frontier for panel data is introduced. Subject unobservable heterogeneity is modeled by a random component and the presence of both a time invariant and a time dependent random inefficiency component is allowed. In this way, subject heterogeneity is not wrongly modeled as inefficiency (Greene, 2005) and it is possible to disentangle a time persistent component from the total inefficiency. The models here presented are related but not identical to the stochastic frontiers introduced by Dominguez-Molina

Roberto Colombi
Dipartimento di Ingegneria dell'Informazione e Metodi Matematici, viale Marconi 5
24044 Dalmine, e-mail: roberto.colombi@unibg.it

et al. (2004) and to the linear mixed models with skew normal errors of Lin and Lee (2005) and Arellano-Valle *et al.* (2005).

2 A Stochastic Frontier for Panel Data

We consider the stochastic frontier model:

$$y_{it} = \beta_0 + \mathbf{x}'_{it}\boldsymbol{\beta} + b_i - u_{i0} - u_{it} + e_{it} \quad (1)$$

where the index i , $i = 1, 2, \dots, n$, denotes n units or subjects and t , $t = 1, 2, \dots, T$, the T time points at which every subject is observed. The variable y_{it} is the logarithm of the output of the i -th unit at time t , $\mathbf{x}'_{it}$ is a row vector of p regressors and $\boldsymbol{\beta}$ is a column vector of unknown parameters. The random variable e_{it} is the *idiosyncratic random component* and b_i is the *random subject effect*. Finally u_{i0} is the *time invariant stochastic inefficiency* and u_{it} is the *time dependent stochastic inefficiency*. We note that $y_{it}^s = \beta_0 + \mathbf{x}'_{it}\boldsymbol{\beta} + b_i + e_{it}$ is the subject specific stochastic frontier and that $y_{it}^p = \beta_0 + \mathbf{x}'_{it}\boldsymbol{\beta} + e_{it}$ is the population stochastic frontier. We assume that: (A1a) for $i=1, 2, \dots, n$, the $2 \times (T+1)$ components u_{i0} , b_i , u_{it} , e_{it} , $t = 1, 2, \dots, T$, of the vector $\boldsymbol{\epsilon}_i$ are independent random variables; (A1b) the random vectors $\boldsymbol{\epsilon}_i$, $i = 1, 2, \dots, n$ are independent; (A2) for every i , u_{i0} is a normal random variable with null expected value and variance σ_{1u}^2 left truncated at zero and b_i is a normal random variable with null expected value and variance σ_b^2 ; (A3) for every i and t , u_{it} is a normal random variable with null expected value and variance σ_{2u}^2 left truncated at zero and e_{it} is a normal random variable with null expected value and variance σ_e^2 ; (A4) the $\mathbf{x}_{it}$ are vectors of known constants. The following special cases of the full model (1) (denoted by TTT) are important:

the model TFT without time dependent inefficiency:

$$y_{it} = \beta_0 + \mathbf{x}'_{it}\boldsymbol{\beta} + b_i - u_{i0} + e_{it} \quad (2)$$

the model FTT without subject specific component:

$$y_{it} = \beta_0 + \mathbf{x}'_{it}\boldsymbol{\beta} - u_{i0} - u_{it} + e_{it} \quad (3)$$

the Greene (2005) true random effect stochastic frontier TTF without time independent inefficiency:

$$y_{it} = \beta_0 + \mathbf{x}'_{it}\boldsymbol{\beta} + b_i - u_{it} + e_{it} \quad (4)$$

the model FFT (Pitt and Lee model I, 1981) without time dependent inefficiency and subject specific component:

$$y_{it} = \beta_0 + \mathbf{x}'_{it}\boldsymbol{\beta} - u_{i0} + e_{it} \quad (5)$$

the pooled stochastic frontier FTF (Pitt and Lee model II, 1981) without time independent inefficiency and subject specific component:

$$y_{it} = \beta_0 + \mathbf{x}'_{it}\boldsymbol{\beta} - u_{it} + e_{it}. \quad (6)$$

The main contribution of this work is that time independent inefficiency can coexist with time dependent inefficiency or subject random effects (as in the new models TTT, FTT, TFT) and that it is not necessary to resort to simulated maximum likelihood methods as suggested by Greene (2005) with respect to the model TTF. Testing one of the previous models against the general model TTT is a non standard problem because under the null hypothesis one or two parameters are on the boundary of the parametric space. In fact, under reasonable assumptions, in this case the asymptotic distribution of the log-likelihood ratio test statistic is a mixture of chi squared distributions known as *chi-bar-squared distribution* (Silvapulle and Sen, 2005). For the sake of simplicity we consider only the case of balanced panel data with a fixed number of observations per unit but the results of the following sections are easily extended to unbalanced panels.

2.1 Log Likelihood of the Closed Skew Normal Stochastic Frontier

Let $\mathbf{1}_T$ be the vector of ones, $\mathbf{0}_T$ be the vector of zeros and $\mathbf{I}_T$ be the identity matrix of dimension T . Moreover $\mathbf{y}_i$ is the vector of the T observations on the i -th unit, $\mathbf{X}_i$ is the $T \times p$ matrix with rows $\mathbf{x}'_{it}$, $\mathbf{u}_i$ is the $(T+1)$ vector with components $u_{i0}, u_{i1}, u_{i2}, \dots, u_{iT}$ and $\mathbf{e}_i$ is the vector of the idiosyncratic random components of the i -th unit. From (1) it follows that: $\mathbf{y}_i = \mathbf{1}_T(\beta_0 + b_i) + \mathbf{X}_i\boldsymbol{\beta} + \mathbf{A}\mathbf{u}_i + \mathbf{e}_i$ where the matrix $\mathbf{A}$ is so defined: $\mathbf{A} = -[\mathbf{1}_T \mathbf{I}_T]$. We also need the following variance matrices:

$$\mathbf{V} = \begin{bmatrix} \sigma_{1u}^2 & \mathbf{0}'_T \\ \mathbf{0}_T & \sigma_{2u}^2 \mathbf{I}_T \end{bmatrix}, \quad \boldsymbol{\Sigma} = \sigma_e^2 \mathbf{I}_T + \sigma_b^2 \mathbf{1}_T \mathbf{1}'_T \quad (7)$$

$$\mathbf{A} = \mathbf{V} - \mathbf{V}\mathbf{A}'(\boldsymbol{\Sigma} + \mathbf{A}\mathbf{V}\mathbf{A}')^{-1}\mathbf{A}\mathbf{V}, \quad \mathbf{R} = \mathbf{V}\mathbf{A}'(\boldsymbol{\Sigma} + \mathbf{A}\mathbf{V}\mathbf{A}')^{-1}. \quad (8)$$

With $\phi_q(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Omega})$ we denote the density function of a q -dimensional normal random variable with expected value $\boldsymbol{\mu}$ and variance $\boldsymbol{\Omega}$ and $\bar{\Phi}_q(\boldsymbol{\mu}, \boldsymbol{\Omega})$ is the probability that a q -variate normal random variable of expected value $\boldsymbol{\mu}$ and variance matrix $\boldsymbol{\Omega}$ belongs to the positive orthant. From now on it will be $q = T+1$.

Proposition 1. *Under the assumptions A1a, A2-A4, the random vector $\mathbf{y}_i$ has a (T, q) closed skew normal distribution with density*

$$f(\mathbf{y}_i) = \phi_T(\mathbf{y}_i, \mathbf{1}_T\beta_0 + \mathbf{X}_i\boldsymbol{\beta}, \boldsymbol{\Sigma} + \mathbf{A}\mathbf{V}\mathbf{A}') \frac{\bar{\Phi}_q(\mathbf{R}(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta} - \mathbf{1}_T\beta_0), \mathbf{A})}{2^{-q}}$$

and moment generating function

$$E(\exp\{\mathbf{t}'\mathbf{y}_i\}) = \frac{\bar{\Phi}_q(\mathbf{V}\mathbf{A}'\mathbf{t}, \mathbf{V})}{2^{-q}} \exp\{\mathbf{t}'(\mathbf{1}_T\beta_0 + \mathbf{X}_i\boldsymbol{\beta}) + \frac{1}{2}\mathbf{t}'(\boldsymbol{\Sigma} + \mathbf{A}\mathbf{V}\mathbf{A}')\mathbf{t}\}.$$

When $\mathbf{y}_i$ is the vector of the logarithms of the outputs and $\mathbf{t}$ is the k -th column of the identity matrix of dimension T , the moment generating function of the previous Proposition gives the expected value of the k -th component of the vector $\exp\{\mathbf{y}_i\}$ of the outputs.

Proposition 2. *Under the assumptions A1-A4, the log-likelihood of nT observations from model (1) is:*

$$L = \sum_{i=1}^n (\ln \phi_T(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta}, \mathbf{1}_T\beta_0, \boldsymbol{\Sigma} + \mathbf{A}\mathbf{V}\mathbf{A}') + \ln \bar{\Phi}_q(\mathbf{R}(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta} - \mathbf{1}_T\beta_0), \mathbf{A})) \quad (9)$$

which is the log-likelihood of the n independent closed skew normal random variables $\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta}$.

The proof of Proposition 1 is given in the Appendix. Proposition 2 is an immediate consequence of assumption A1b and Proposition 1. Note that the computational complexity of the evaluation of the log-likelihood (9) comes from the multiple integrals $\bar{\Phi}_q(\mathbf{R}(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta} - \mathbf{1}_T\beta_0), \mathbf{A})$, $i = 1, 2, \dots, n$. There are however remarkable exceptions. For the Pitt and Lee models TFT, FFT with and without the subject random effect, the above integral is one dimensional and for the pooled model FTF it is a product of T one dimensional integrals. Thus, the computational problem is limited to the models where the time dependent inefficiency is present together with the subject specific component or the time invariant inefficiency. For a survey on efficient numerical and Monte Carlo methods to compute the previous multinormal integral see Genz and Bretz (2009).

3 A Two Step Estimation Method

Instead of using the *maximum likelihood estimator* that maximizes the log-likelihood (9) we propose the following two step or *pseudo maximum likelihood estimator* (Parke, 1986). The idea behind the first step of our proposal is that consistent estimators of the components of $\boldsymbol{\beta}$ are given by standard methods of estimation used in the context of panel data or of mixed models. After having estimated $\boldsymbol{\beta}$, one is interested in assessing the relevance of the inefficiency random components and of the subject specific effect, which requires the estimation of the variances σ_{1u}^2 , σ_{2u}^2 , σ_b^2 , σ_e^2 . This task is performed by

the second step of our procedure. In our experience the two step procedure is less time consuming than the maximum likelihood method and more stable with respect to bad starting values. Moreover, the two step estimates are good starting point for the search of the maximum likelihood estimates. The two step procedure is particularly convenient when different sets of regressors must be tried in order to specify the deterministic part of the frontier. According to our procedure this can be done without worrying about σ_{1u}^2 , σ_{2u}^2 , σ_b^2 , σ_e^2 . Let s_i be the subject specific random components $s_i = b_i - u_{i0}$ and e_{it}^* be the random components $e_{it}^* = -u_{it} + e_{it}$ (the variances and expected values of the previous random variables can be easily computed but we omit the needless details). In the first step we compute the estimate $\tilde{\beta}$ of β by a standard estimation method for panel data models or mixed models applied to the random effects model: $y_{it} = \beta_0 + \mathbf{x}_{it}'\beta + s_i + e_{it}^*$. From A.1-A.4 it follows that $\tilde{\beta}$ is a consistent estimator of β (β_0 is not consistently estimated because the random components of the previous model have non-null expected value). In the second step, we use the fact that $\mathbf{y}_i - \mathbf{X}_i\beta$, $i = 1, 2, \dots, n$ are independent closed skew normal random variables with log-likelihood given by (9) and we maximize the pseudo log-likelihood

$$\tilde{L} = \sum_{i=1}^n (\ln \phi_T(\mathbf{y}_i - \mathbf{X}_i\tilde{\beta}, \mathbf{1}_T\beta_0, \Sigma + \mathbf{A}\mathbf{V}\mathbf{A}) + \ln \bar{\Phi}_q(\mathbf{R}(\mathbf{y}_i - \mathbf{X}_i\tilde{\beta} - \mathbf{1}_T\beta_0))) \quad (10)$$

with respect to β_0 , σ_b^2 , σ_{1u}^2 , σ_{2u}^2 , σ_e^2 .

The properties of these estimators follows from the general theory of two step M-estimators (Parke 1986)

4 Prediction of the Random Components

When $\mathbf{y}_i$ is the vector of the logarithms of the outputs, an important topic in applied research is forecasting the subject random components by the expected values $E(\exp\{b_i\}|\mathbf{y}_i)$ and the subject random inefficiencies by $E(\exp\{-\mathbf{u}_i\}|\mathbf{y}_i)$. This section shows how the relevant expected values can be computed. It is convenient to introduce the following definitions:

$$\mathbf{r}_i = \mathbf{y}_i - \mathbf{X}_i\beta - \mathbf{1}_T\beta_0, \quad \tilde{\sigma}_b^2 = \sigma_b^2 - \sigma_b^4 \mathbf{1}_T' \Psi \mathbf{1}_T \quad (11)$$

$$\Psi = (\Sigma + \mathbf{A}\mathbf{V}\mathbf{A}')^{-1}, \quad \tilde{\Lambda} = \Lambda - \mathbf{R}\mathbf{1}_T\mathbf{1}_T'\mathbf{R}' \frac{\sigma_b^4}{\tilde{\sigma}_b^2}. \quad (12)$$

Proposition 3. *a) Conditionally on $\mathbf{y}_i$ the subject specific random component b_i has a $(1, q)$ closed skew normal distribution with density*

$$f(b_i|\mathbf{y}_i) = \phi(b_i, \sigma_b^2 \mathbf{1}_T' \Psi \mathbf{r}_i, \tilde{\sigma}_b^2) \frac{\bar{\Phi}_q(\mathbf{R}\mathbf{r}_i - \mathbf{R}\mathbf{1}_T\sigma_b^2\tilde{\sigma}_b^{-2}(b_i - \sigma_b^2 \mathbf{1}_T' \Psi \mathbf{r}_i), \tilde{\Lambda})}{\bar{\Phi}_q(\mathbf{R}\mathbf{r}_i, \Lambda)}$$

b) conditionally on $\mathbf{y}_i$ the random inefficiency vector $\mathbf{u}_i$ is a left truncated normal random variable with density:

$$f(\mathbf{u}_i|\mathbf{y}_i) = \frac{\phi_q(\mathbf{u}_i, \mathbf{R}\mathbf{r}_i, \mathbf{\Lambda})}{\bar{\Phi}_q(\mathbf{R}\mathbf{r}_i, \mathbf{\Lambda})}, \quad \mathbf{u}_i > \mathbf{0}$$

c)

$$E(\exp\{b_i\}|\mathbf{y}_i) = \frac{\bar{\Phi}_q(\mathbf{R}\mathbf{r}_i - \mathbf{R}\mathbf{1}_T\sigma_b^2, \mathbf{\Lambda})}{\bar{\Phi}_q(\mathbf{R}\mathbf{r}_i, \mathbf{\Lambda})} \exp\{\sigma_b^2\mathbf{1}_T'\boldsymbol{\Psi}\mathbf{r}_i + \frac{1}{2}\tilde{\sigma}_b^2\}$$

d)

$$E(\exp\{\mathbf{t}'\mathbf{u}_i\}|\mathbf{y}_i) = \frac{\bar{\Phi}_q(\mathbf{R}\mathbf{r}_i + \mathbf{\Lambda}\mathbf{t}, \mathbf{\Lambda})}{\bar{\Phi}_q(\mathbf{R}\mathbf{r}_i, \mathbf{\Lambda})} \exp\{\mathbf{t}'\mathbf{R}\mathbf{r}_i + \frac{1}{2}\mathbf{t}'\mathbf{\Lambda}\mathbf{t}\}$$

The proof of the Proposition is given in the Appendix. If $-\mathbf{t}'$ is the k -th row of the identity matrix of dimension q the result d) of the previous Proposition gives the conditional expected value of the k -th component of the inefficiency vector $\exp\{-\mathbf{u}_i\}$. In particular for $k = 1$ we get the conditional expected value of the time invariant inefficiency. We stress also that, conditionally on the observation $\mathbf{y}_i$, the random effect has not a normal distribution as it happens in standard random effect models.

5 Example

We consider the Rice panel data set examined by Battese *et al.* (2005, pp. 325-326). In this data set the *annual rice production* (in tonnes) of $n = 43$ rice producers in the Tarlac region of the Philippines is reported from 1990 to 1997. The Rice production is the output and the inputs are the *area planted* (hectares), the *labor* used (man days) and the *fertiliser* used (kg). In this application the log-linear stochastic frontier is intended to model the maximum rice production in function of the inputs. The random component b_i models unobserved differences among producers and the inefficiency components model the producer inability to reach the optimal production. The results reported in the Table 1 were obtained (only the maximum of the log-likelihood and the estimated variances are reported).

It clearly emerges from Table 1 that both the time invariant and time dependent inefficiency component are relevant and that their omission inflates the variance of the idiosyncratic random component. On the contrary the subject specific random component seem not to be relevant. The only models to be considered are obviously the models TTT and FTT. The asymptotic distribution of the log-likelihood ratio statistics to test FTT against TTT in this case is a 0.5 mixture of a chi squared distribution with zero degrees

Table 1 Skew normal sthochastic frontiers for the Rice data

model ¹	max. log-lik.	σ_b^2	σ_{2u}^2	σ_{1u}^2	σ_e^2
TTT	-65.340	0.0001	0.1744	0.0661	0.0161
TFT	-86.430	0	0	0.0721	0.0832
FTT	-65.371	0	0.1778	0.650	0.0159
TTF	-68.117	0.0228	0.1794	0	0.0154
FFT	-86.431	0	0	0.0723	0.0832
FTF	-104.907	0	0	0	0.1077
TFF ²	-88.605	0.1584	0	0	0.2901
FFF ²	-104.907	0	0	0	0.3283

¹ The R (R Development Core Team, 2009) functions: SNF_twostage and SNF_maxlik, developed by the author, were used to fit the models of the Table.

² TFF and FFF denote the subject random effect regression model and the classical regression model respectively.

of freedom and a chi squared distribution with one degree of freedom. The p-value is P=0.402 and so FTT can be maintained.

Appendix

For an easy reference we report the definition of the closed skew normal distribution (Arellano-Valle and Azzalini, 2006, Gonz  les-Far  as *et al.*, 2004).

Definition 1. A random vector $\mathbf{x}$ has a (p, q) closed skew normal distribution if its probability density function is:

$$f(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Gamma}, \mathbf{D}, \boldsymbol{\nu}, \boldsymbol{\Delta}, p, q) = \frac{\phi_p(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Gamma}) \bar{\Phi}_q(\mathbf{D}(\mathbf{y} - \boldsymbol{\mu}) - \boldsymbol{\nu}, \boldsymbol{\Delta})}{\bar{\Phi}_q(-\boldsymbol{\nu}, \boldsymbol{\Delta} + \mathbf{D}\boldsymbol{\Gamma}\mathbf{D}')}.$$

The moment generating function of the previous random variable is:

$$E(\exp\{\mathbf{t}'\mathbf{x}\}) = \frac{\bar{\Phi}_q(\mathbf{D}\boldsymbol{\Gamma}\mathbf{t} - \boldsymbol{\nu}, \boldsymbol{\Delta} + \mathbf{D}\boldsymbol{\Gamma}\mathbf{D}')}{\bar{\Phi}_q(-\boldsymbol{\nu}, \boldsymbol{\Delta} + \mathbf{D}\boldsymbol{\Gamma}\mathbf{D}')} \exp\{\mathbf{t}'\boldsymbol{\mu} + \frac{1}{2}\mathbf{t}'\boldsymbol{\Gamma}\mathbf{t}\}. \quad (13)$$

Proof of Proposition 1. From the assumptions A1a, A2-A4 it follows that

$$f(\mathbf{y}_i) = \phi_T(\mathbf{y}_i, \mathbf{1}_T\beta_0 + \mathbf{X}_i\boldsymbol{\beta}, \boldsymbol{\Sigma} + \mathbf{A}\mathbf{V}\mathbf{A}') \times \frac{\int_0^\infty \dots \int_0^\infty \phi_q(\mathbf{u}, \mathbf{R}(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta} - \mathbf{1}_T\beta_0), \mathbf{A}) du_0 \dots du_T}{\int_0^\infty \dots \int_0^\infty \phi_q(\mathbf{u}, \mathbf{0}, \mathbf{V}) du_0 \dots du_T} \quad (14)$$

which proves the first part of the Proposition. It is also straightforward to see that $f(\mathbf{y}_i)$ is an instance of the closed skew normal distribution with: $p = T$, $\boldsymbol{\nu} = \mathbf{0}$, $\boldsymbol{\Delta} = \mathbf{A}$, $\mathbf{D} = \mathbf{R}$, $\boldsymbol{\Gamma} = \boldsymbol{\Sigma} + \mathbf{A}\mathbf{V}\mathbf{A}'$. Thus, the second part of the Proposition follows from (13).

Proof of Proposition 3. To prove b) note that:

$$f(\mathbf{u}_i|\mathbf{y}_i) = \frac{\phi_T(\mathbf{y}_i, \mathbf{1}_T\beta_0 + \mathbf{X}_i\beta, \Sigma + \mathbf{A}\mathbf{V}\mathbf{A}')\phi_q(\mathbf{u}, \mathbf{R}\mathbf{r}_i, \Lambda)2^q}{f(\mathbf{y}_i)} \quad (15)$$

and use the result of Proposition 1 about $f(\mathbf{y}_i)$. Noting that $f(\mathbf{u}_i|\mathbf{y}_i)$ is the density of a left truncated at zero multi-normal random variable, d) follows from Lemma 13.6.1 of Dominguez-Molina, et al. (2004). To prove a) we observe that:

$$\begin{aligned} f(b_i|\mathbf{y}_i) = & \phi(b_i, \sigma_b^2 \mathbf{1}'\Psi\mathbf{r}_i, \tilde{\sigma}_b^2) \times \\ & \times \frac{\int_0^\infty \dots \int_0^\infty \phi_q(u, \mathbf{R}\mathbf{r}_i - \mathbf{R}\mathbf{1}_T\sigma_b^2\tilde{\sigma}_b^{-2}\tilde{b}_i, \tilde{\Lambda})du_0\dots du_T}{\int_0^\infty \dots \int_0^\infty \phi_q(\mathbf{u}, \mathbf{R}\mathbf{r}_i, \Lambda)du_0\dots du_T} \end{aligned} \quad (16)$$

Point a) of the proposition follows immediately. Observing that $f(b_i|\mathbf{y}_i)$ is an instance of the closed skew normal density, point c) follows from the result (13) on the moment generating function of a closed skew normal random variable.

References

1. Arellano-Valle, E., Bolfarine, H., Lachos, H.: Skew-normal linear mixed models. *J. Data. Sci.* **3**, 415-438 (2005)
2. Arellano-Valle, E., Azzalini A.: On the unification of families of skew-normal distributions. *Scand. J. Stat.* **33**, 561-574 (2006)
3. Battese, G., Coelli, T., O'Donnel, C., Rao, D.: An introduction to Efficiency and Productivity Analysis. Springer, New York (2005)
4. Dominguez-Molina, A., Gonzales-Farias, G., Ramos-Quiroga, R.: Skew normality in stochastic frontier analysis. In: Genton, M. (ed) *Skew Elliptical Distributions and their Applications*, pp. 223-241. Chapman and Hall CRC, London (2004)
5. Gonz  les-Far  as, G., Dom  nguez Molina, A., Gupta, A.: Additive properties of skew normal random vectors. *J. Stat. Plan. Infer.* **126**, 521-534 (2004)
6. Genz, A., Bretz, F.: *Computation of Multivariate Normal and t Probabilities*. Springer, New York (2009)
7. Greene, W.: Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. *J.Econometrics.* **126**, 269-303 (2005)
8. Kumbhakar, S., Lovell K.: *Stochastic Frontier Analysis*. Cambridge University Press, Cambridge (2000)
9. Lin, T., Lee, C.: Estimation and prediction in linear mixed models with skew-normal random effects for longitudinal data. *Statist.Med.* **27**, 1490-1507 (2005)
10. Parke, W.: Pseudo maximum likelihood estimation: the asymptotic distribution. *Ann.Stat.* **14**, 355-357 (1986).
11. Pitt, M., Lee, L.: Measurement of sources of technical inefficiency in the Indonesian weaving industry. *J.Dev.Econ.* **9**, 43-64 (1981)
12. R Development Core Team: *R: A language and environment for statistical computing*. R Foundation for Statistical Computing ISBN 3-900051-07-0 (2009) <http://www.R-project.org>
13. Sivapulle, M. J., Sen, P. K.: *Constrained Statistical Inference*. Wiley, Hoboken (2005)